Wolf, B. A., Colca, J. R., Turk, J., Florholmen, J., & McDaniel, M. L. (1988a) *Am. J. Physiol. 254*, E121–E136.
Wolf, B. A., Florholmen, J., Turk, J., & McDaniel, M. L. (1988b) *J. Biol. Chem. 263*, 3565–3575.
Wollheim, C. B., & Sharp, G. W. G. (1981) *Physiol. Rev. 61*, 914–973.
Wollheim, C. B., & Biden, T. J. (1986) *J. Biol. Chem. 261*, 8314–8319.
Wollheim, C. B., Dunne, M. J., Peter-Riesch, B., Bruzzone, R., Pozzan, T., & Petersen, O. H. (1988) *EMBO J. 7*, 2443–2449.
Zawalich, W. S. (1988) *Diabetalogia 31*, 435–442.
Zawalich, W., Brown, C., & Rasmussen, H. (1983) *Biochem. Biophys. Res. Commun. 117*, 448–455.
Zawalich, W. S., Diaz, V. A., & Zawalich, K. C. (1988) *Am. J. Physiol. 254*, E609–E626.

Conformation of Recombinant Desulfatohirudin in Aqueous Solution Determined by Nuclear Magnetic Resonance†

Hideyuki Haruyama‡ and Kurt Wüthrich*

Institut für Molekularbiologie und Biophysik, Eidgenössiche Technische Hochschule—Hönggerberg, CH-8093 Zürich, Switzerland

Received October 7, 1988; Revised Manuscript Received January 3, 1989

ABSTRACT: The three-dimensional structure of recombinant desulfatohirudin in aqueous solution was determined by ^{1}H nuclear magnetic resonance at 600 MHz and distance geometry calculations with the program DISMAN. The input for the structure calculations was prepared on the basis of complete sequence-specific resonance assignments at pH 4.5 and 22 °C and consisted of 425 distance constraints from nuclear Overhauser enhancements and 159 supplementary constraints from spin–spin coupling constants and from the identification of intramolecular hydrogen bonds. Residues 3–30 and 37–48 form a molecular core with two antiparallel β-sheets and several well-defined turns. The three disulfide bonds 6–14, 16–28, and 22–39 were identified by NMR. In contrast to this well-defined molecular core, with an average root mean square distance for the polypeptide backbone of 0.8 Å for a group of nine DISMAN solutions, no preferred conformation was found for the C-terminal segment 49–65, and a loop consisting of residues 31–36 is not uniquely constrained by the NMR data either. These structural properties of recombinant desulfatohirudin coincide closely with the previously described solution conformation of natural hirudin, but the presence of localized differences is indicated by chemical shift differences for residues Asp 5, Ser 9, Leu 15, Asp 53, Gly 54, and Asp 55.

Hirudin is a small protein of 65 amino acid residues, which was first isolated from the salivary glands of the leech *Hirud medicinalis* (Markwardt, 1970). Hirudin acts as a specific inhibitor of α-thrombin. Strong and exclusive affinity for α-thrombin, low antigenicity, and rapid clearance from the blood make it attractive for medical applications (Markwardt et al., 1982). While its fundamental functional mechanisms were extensively investigated (Chang, 1983; Stone & Hofsteenge, 1986), insufficient amounts of homogeneous hirudin preparations could be obtained from natural sources to warrant practical medical applications. With recombinant methods large amounts of the protein can now be produced (Meyhack et al., 1987; Grossenbacher et al., 1987), which also renews interest in the structure–function correlations of this protein (Braun et al., 1988).

Natural hirudin contains a sulfated tyrosine residue in position 63 (Badgy et al., 1976; Dodt et al., 1984). The recombinant protein lacks this posttranscriptional modification and contains normal tyrosine in position 63. At the same time its inhibitory activity toward α-thrombin is 10 times weaker than that of native hirudin. However, the reduced activity of the recombinant protein cannot at present be simply correlated with the absence of sulfated Tyr 63, since neither the location of the disulfide bridges in desulfatohirudin nor its three-dimensional structure is known. Similar three-dimensional structures for the two proteins were indicated by the observation that the activity of recombinant hirudin corresponds approximately to that of chemically desulfated hirudin (Braun et al., 1988). To establish more direct correlations between the reduced activity of the recombinant desulfatohirudin and its structure, this paper describes the determination of the three-dimensional desulfatohirudin structure in solution by NMR.[1] This structure is then compared with the previously reported solution conformation of natural hirudin (Clore et al., 1987; Sukumaran et al., 1987).

† Financial support for this project was obtained from Sankyo Co., Ltd., Tokyo, Japan.

‡ Present address: Analytical and Metabolic Research Laboratories, Sankyo Co., Ltd., 2-58 Hiromachi 1-chome, Shinagawa-ku, Tokyo, 140 Japan.

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D, two dimensional; 2QF-COSY, two-quantum-filtered homonuclear correlated spectroscopy; E-COSY, two-dimensional exclusive correlation spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; TOCSY, two-dimensional total correlation spectroscopy; TPPI, time-proportional phase incrementation; RELAYED-COSY, two-dimensional relayed coherence-transfer spectroscopy; TSP, 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate, sodium salt; RMSD, root mean square difference.

0006-2960/89/0428-4301$01.50/0

Materials and Methods

Recombinant desulfatohirudin was obtained from Ciba-Geigy AG., Basel, Switzerland (Meyhack et al., 1987; Grossenbacher et al., 1987). For the NMR experiments 6.5 mM solutions were prepared in ^2H_2O or in a mixed solvent of 90% H_2O/10% ^2H_2O. For the identification of the slowly exchanging amide protons, a freshly prepared solution of the protein in ^2H_2O was studied. For further measurements in ^2H_2O, all labile protons were replaced with deuterium by keeping the sample at 50 °C for 20 min followed by repeated lyophilization from ^2H_2O. The sequence-specific assignments were obtained primarily from a set of spectra recorded at 22 °C and pH 4.5, where the pH value was adjusted by the addition of dilute HCl/NaOH or ^{2}HCl/NaO^{2}H, respectively. Additional experiments in H_2O solution were recorded at 15 °C and pH 4.5 and at 22 °C and pH 3.0, respectively, to resolve overlap of amide proton chemical shifts.

Proton NMR spectra were recorded on Bruker AM 600 and WH 500 spectrometers in the pure phase absorption mode with time proportional phase incrementation (Redfield & Kunz, 1975; Marion & Wüthrich, 1983). Standard procedures were used for the following experiments, which were needed for the resonance assignments and for the preparation of the input for the structure calculations: 2QF-COSY (Rance et al., 1984; Neuhaus et al., 1985), TOCSY with a mixing time τ_m of 100 ms (Braunschweiler & Ernst, 1983; Bax & Davis, 1985), NOESY with τ_m = 40, 70, and 100 ms (Jeener et al., 1979; Anil Kumar et al., 1980), RELAYED-COSY with τ_m = 38 ms (Wagner, 1983), and two-quantum spectroscopy (Wagner & Zuiderweg, 1983; Otting & Wüthrich, 1986). The typical data size was 2048 × 512 points. The final digital resolution after zero filling was 2.7 Hz/point along ω_2 and 5.4 Hz/point along ω_1. For the NOESY spectra in ^2H_2O, systematic incrementation of the mixing time was used to minimize the contributions from zero-quantum coherence (Rance et al., 1985). To improve the signal-to-noise ratio of weak cross peaks in the NOESY spectra, a cosine square window was applied (Kline et al., 1988). Base-line distortions resulting from this window function were eliminated by a base-line correction with a third- or fourth-order polynomial. For the measurements of the spin–spin coupling constants $^3J_{HN\alpha}$, a 2QF-COSY spectrum in H_2O was recorded with a final digital resolution of 0.35 Hz/point along ω_2. The spin-spin coupling constants $^3J_{\alpha\beta}$ were measured in a phase-sensitive E-COSY spectrum (Griesinger et al., 1985).

Stereospecific assignments for prochiral groups of protons were obtained with the program HABAS (Güntert et al., 1989). This program makes similar use of the spin–spin coupling constants $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ and of intraresidual and sequential NOEs as in previously described manual approaches (Arseniev et al., 1988; Hyberts et al., 1987; Kline et al., 1988). A few additional stereospecific assignments were obtained from analysis of long-range NOEs (Senn et al., 1984; Zuiderweg et al., 1985; Kline et al., 1988).

To relate the NOE data with corresponding ^{1}H–^{1}H distances, three NOESY spectra with mixing times of 40, 70, and 100 ms were recorded. The relative peak intensities at τ_m = 70 ms were evaluated by counting the number of contour levels at the peak maxima, and the other spectra were used for controls. We then largely followed earlier practice (Williamson et al., 1985; Kline et al., 1988; Wüthrich, 1986): The intraresidual and sequential NOEs were divided into three classes with different upper bounds, by use of calibration curves based on cross peaks which correspond to known distances in the regular secondary structures (Wüthrich et al., 1984): $d_{\alpha N}$ (β-sheet) = 2.2 Å; $d_{N\alpha}(i,i)$ (β-sheet) = 2.8 Å; $d_{\beta\beta}(i,i)$ = 1.8 Å; $d_{\alpha\alpha}(i,j)$ between opposite residues in neighboring antiparallel β-strands = 2.3 Å. For $d_{\alpha N}$ and $d_{N\alpha}(i,i)$, we then used upper bounds of 2.5, 2.8, and 3.2 Å. For the remaining intraresidual and sequential constraints, the corresponding limits were 2.5, 3.0, and 4.0 Å. With few exceptions the upper limit for medium-range and long-range backbone proton constraints was 4.0 Å, and for interresidue NOEs with side chain protons it was 5.0 Å, independent of the peak intensity. Exceptions were made for approximately 10 well-isolated cross peaks with outstandingly strong intensities, where these upper bounds were lowered to 3.0 and 4.0 Å, respectively.

Supplementary conformational constraints in the form of allowed ranges for the dihedral angles were obtained as a side product of the elucidation of stereo-specific assignments with the program HABAS. For hydrogen bonds C=O···H—N identified during a preliminary secondary structure analysis (see Figure 5), the O–H distance was constrained between 1.8 and 2.0 Å and the N–O distance between 2.7 and 3.0 Å. For the disulfide bonds the S_i^γ–S_j^γ distance was constrained in the range from 2.0 to 2.1 Å, and the C_i^β–S_j^γ or S_i^γ–C_j^β distance from 3.0 to 3.1 Å (Williamson et al., 1985).

Three-dimensional molecular structures were calculated from the NMR constraints with the distance geometry program DISMAN (Braun & Gō, 1985). The calculations first yielded a group of four structures with an average RMSD for the backbone atoms of nearly 3.0 Å. These low-resolution structures were used to resolve ambiguities in the assignment of NOESY cross peaks (Wüthrich, 1986) and thus to obtain a more extensive data set for a next cycle of structure calculations. After each cycle only those cross peaks which satisfied the following criteria were accepted as additional input data: First, the assignment had to be supported by all the structures calculated from the preliminary set of input data. Second, the peaks produced by the same spin must all have the same line shape. Third, a cross peak was accepted only if it could be identified also in the NOESY spectra recorded with a mixing time of 40 ms, so that cross peaks generated by spin-diffusion were eliminated. The preliminary DISMAN calculations also showed that the peptide segments Thr 4 to Leu 15 and Ser 32 to Lys 36 form loops for which convergence was not as good as for the rest of the molecule. Following a previously used strategy for improving local convergence of DISMAN calculations (Kline et al., 1988), 200 initial structures were generated for these segments and minimized relative to the local constraints. For the final calculations of the whole structure, the starting values for the individual dihedral angles were then randomly selected within the ranges covered by the converged structures for these isolated loops, rather than over the entire accessible range. These supplementary dihedral angle constraints were kept in the data file only until the level of the target function exceeded the length of the peptide segment for which this pretreatment was performed.

Structure calculations were done both for the polypeptide segment 1–49 and for the complete recombinant desulfatohirudin molecule 1–65. In each case, 50 starting structures were generated. Those structures with target functions of less than 5.0 at target level 5 were subjected to further calculations at levels 15, 30, 49, and 65. At each level, 500 cycles of minimization were performed. At every stage those structures which had no individual constraint violations greater than 1.0 Å and no significant violations of van der Waals constraints were selected for the next target-level calculation. Finally, the structures with a target function value of less than 10 were subjected to further minimizations with 1000 iterations at levels

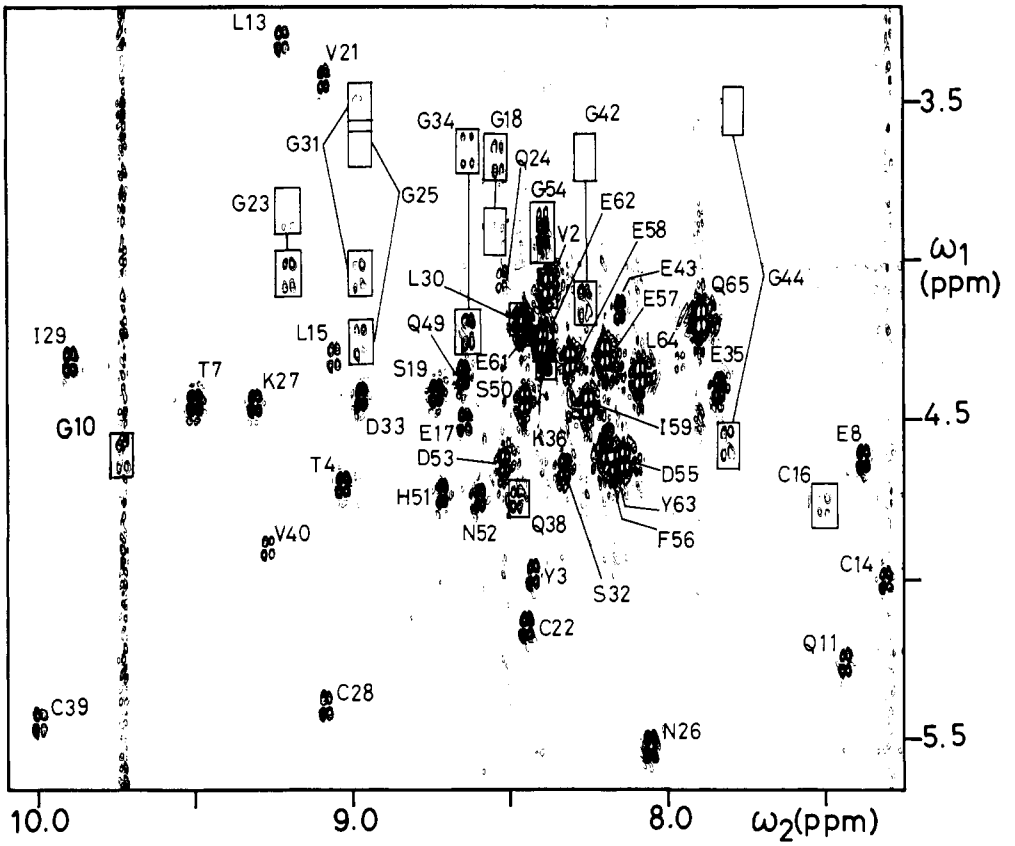

FIGURE 1: Fingerprint region of a 2QF-COSY spectrum of a 6.5 mM solution of recombinant desulfatohirudin in H_2O at 22 °C, pH 4.5. The individual cross peaks are labeled with the amino acid type and the sequence number obtained as the result of the sequential assignments.

5, 15, 30, and the whole sequence of 49 or 65 residues, respectively. After this treatment the structures with a target function value of less than 5 were accepted as converged structures.

RESULTS

Sequence-Specific Resonance Assignments. Using a standard strategy centered around the sequential assignment technique (Billeter et al., 1982; Wagner & Wüthrich, 1982; Wider et al., 1982; Wüthrich, 1986), complete sequence-specific ^1H NMR assignments were obtained at pH 4.5 and 22 °C, and nearly complete sequence-specific assignments of the backbone protons were obtained at pH 3.0 as well. The chemical shifts are listed in Table I, and Figure 1 affords a view of the fingerprint region from a 2QF-COSY spectrum with sequence-specific resonance assignments. The assignments at pH 4.5 provided the basis for the determination of the three-dimensional structure described in this paper. The additional data at pH 3.0 were needed for a direct comparison of recombinant desulfatohirudin with natural hirudin (see Table I), and for studies on dynamic aspects of the desulfatohirudin conformation (Haruyama et al., 1989).

Stereospecific ^1H NMR Assignments. The polypeptide chain of hirudin contains 72 prochiral pairs of methylene protons and methyl groups. For 30 of these, stereospecific assignments could be obtained. Table II lists the method by which the stereospecific assignments were determined in the individual cases.

For the β-methylene protons indicated by the letter H (Table II), the stereospecific assignments were based on an analysis of intraresidual and sequential NOEs and on the vicinal coupling constants $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ with the program HABAS (Güntert et al., 1989). Of the 46 βCH_2 groups in hirudin, 17 could thus be stereospecifically assigned, which includes those in the three prolines. This is well within the extent that one can expect, on fundamental grounds, to obtain stereospecific assignments from local constraints measured with the precision that could be achieved in the present work with desulfatohirudin (Güntert et al., 1989).

For other types of prochiral groups, e.g., δCH_3 of Leu or γCH_3 of Val, stereospecific assignments may be derived from long-range NOE constraints in the three-dimensional structure (method B in Table II). If a sufficiently large number of long-range NOE constraints with the protons of interest are available as input for the structure calculation, one may find that only one of the two possible stereospecific assignments is compatible with the experiments (Kline et al., 1988). Table III illustrates how this approach was used with Leu 15. The first column identifies protons which interact with the δ-methyl groups. The second column lists the observed relative NOE intensities from a given proton to the two δ-methyl groups. The third column lists the range of differences between the distances from proton X to the two methyls in a group of preliminary DISMAN structures of desulfatohirudin obtained without stereospecific assignments. When the range of these distance differences does not include zero, as is the case for all protons X in Table III, the same δ-methyl group must be closer to the proton X in all structures. The fourth column of Table III indicates which of the two possible stereospecific assignments makes the observed NOE intensities compatible with the corresponding interproton distances. In the case of Leu 15, all NOE connectivities lead to the same stereospecific assignment, which was therefore included in Table II. In all, four assignments in Table II are based on this approach.

For the side-chain amide protons of Asn and Gln (method A in Table II) the stereospecific assignments were obtained from comparison of the relative intensities of the intraresidual NOEs to the two β- or γ-methylene protons, respectively, using

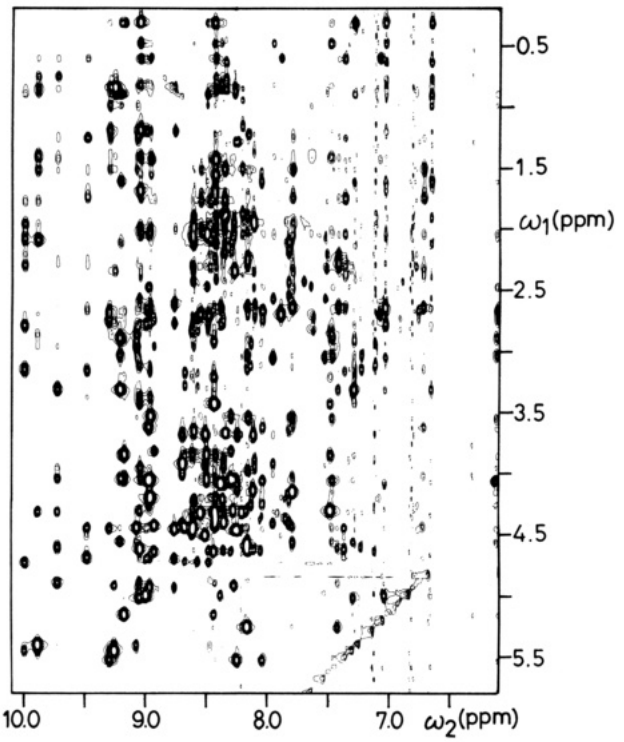

FIGURE 2: Spectral region (ω_1 = 0.2–5.8 ppm, ω_2 = 6.1–10.1 ppm) from a ^{1}H NOESY spectrum at 600 MHz of 6.5 mM recombinant desulfatohirudin in H_2O, pH 4.5, T = 22 °C, recorded with a mixing time of 100 ms.

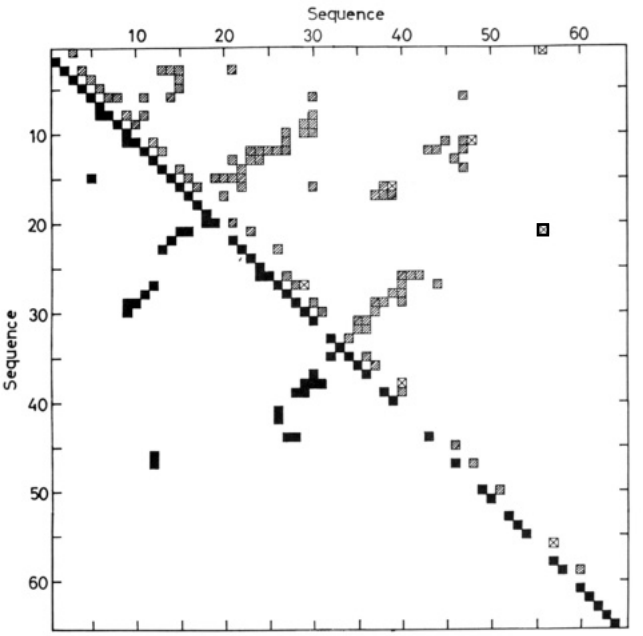

FIGURE 3: Diagonal plot of the NOE distance constraints for recombinant desulfatohirudin. Both axes are calibrated with the sequence of the protein. The squares connect pairs of residues linked by one or several NOE constraints. The filled squares in the lower left part indicate constraints between backbone protons. The hatched squares and crossed squares in the upper right half indicate backbone–side chain and side chain–side chain NOEs, respectively. Whenever a hatched and a crossed square occur at the same place, only the hatched square is shown.

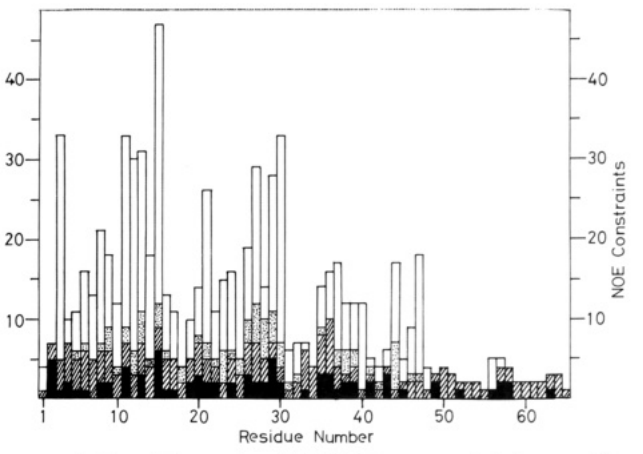

FIGURE 4: Plot of the number of NOE distance constraints per residue versus the amino acid sequence of recombinant desulfatohirudin. All constraints appear twice, once for each of the two interacting residues. Four types of NOE constraints are specified as follows: black, intraresidual; hatched, sequential backbone; dotted; medium-range and long-range backbone; white, interresidual with side-chain protons.

that the geometry of the amide group relative to these protons is fixed.

Preparation of the Input for the Structure Calculations. To illustrate the quality of the data used to collect the NOE distance constraints, Figure 2 shows part of a NOESY spectrum in H_2O. Such experiments were recorded with mixing times of 40, 70, and 100 ms at 14 and 22 °C. In desulfatohirudin the amide proton chemical shifts are sufficiently different at these two temperatures to resolve most ambiguities in cross-peak assignments arising from chemical shift degeneracy. For nonlabile protons such ambiguities were resolved primarily by going through several cycles of resonance assignments and structure calculations. Semiquantitative relations between NOE intensity and ^{1}H–^{1}H distance were used to derive upper bounds on the proton–proton distances. (Details of these procedures are given under Materials and Methods.) The resulting input list of 82 intraresidual constraints, 62 sequential backbone constraints, 37 medium-range and long-range backbone constraints, and 244 interresidual constraints with side-chain protons has been deposited as supplementary material (see paragraph at end of paper regarding supplementary material). Figures 3 and 4 present respectively a diagonal plot of the NOE data and a histogram of the distribution of NOE constraints along the amino acid sequence.

A complete list of the supplementary constraints on the dihedral angles Φ, ψ, and χ^1 obtained with the program HABAS has been deposited as supplementary material.

Hydrogen-bond constraints (see Materials and Methods) included in the input were obtained with the following criteria: By an empirical search for corresponding patterns of conformational constraints, the principal secondary structure elements were identified prior to the distance geometry calculations (Wüthrich et al., 1984; Wüthrich 1986). In desulfatohirudin two antiparallel β-sheets and several tight turns were thus found (Figure 5). In these secondary structures, those hydrogen bonds were selected for inclusion in the input, for which the amide proton exchange was slow and the acceptor group could be unambiguously identified from medium-range or long-range NOEs (arrows in Figure 5). This procedure yielded constraints for 13 hydrogen bonds (asterisks in Table IV).

Identification of the Disulfide Bonds. In natural hirudin the disulfide bridges have been determined by chemical methods as Cys 6 to Cys 14, Cys 16 to Cys 28, and Cys 22 to Cys 39 (Dodt et al., 1985). In recombinant desulfatohirudin the following NMR data showed that this protein contains the same disulfide bonds: Long-range NOEs observed between the backbone protons of Cys 6 and βCH_2 of Cys 14 gave direct evidence for disulfide bridge formation between these two Cys

Table I: ^{1}H Chemical Shifts, δ (ppm), Relative to Internal TSP for Recombinant Desulfatohirudin in H_2O Solution at 22 °C and pH 4.5 and at pH 3.0[a]

amino acid residue	δ, pH 4.5				δ, pH 3.0		
	NH	αH	βH	others	NH	αH	others
Val 1		3.66	1.99	γCH$_3$ 0.64, 0.78		3.64	
Val 2	8.37	4.11	1.89	γCH$_3$ 0.82, 0.87	8.34	4.07	
Tyr 3	8.43	4.99	2.65, 2.79	δH 7.04; εH 6.67	8.43	4.94	
Thr 4	9.02	4.71	4.46	γCH$_3$ 1.21	8.99	4.66	
Asp 5	8.80	4.93	2.62, 2.77		***8.99***	**4.94**	βCH$_2$ **2.76**, **2.87**
Cys 6	8.96	4.72	2.68, 2.77		8.97	4.64	
Thr 7	9.49	4.45	4.51	γCH$_3$ 1.27	9.50	4.44	
Glu 8	7.36	4.63	1.75, 2.05	γCH$_2$ 2.34, 2.41	7.38	4.62	γCH$_2$, 2.45, –
Ser 9	9.09	4.89	3.95, 4.04		**9.07**	3.93	
Gly 10	9.73	3.32, 4.61			9.69	3.28, 4.61	
Gln 11	7.42	5.29	2.26, 2.31	γCH$_2$ 2.63; εNH$_2$ 6.73, 7.82	7.42	*5.18*	
Asn 12	8.16	4.54	2.89, 3.04	δNH$_2$ 6.12, 7.48	8.16	4.53	
Leu 13	9.21	3.32	2.14, 2.14	γH 1.62; δCH$_3$ 0.91, 1.00	9.19	3.29	
Cys 14	7.29	5.01	3.13, 3.39		7.29	4.98	
Leu 15	9.05	4.31	**1.42**, **2.00**	γH **1.20**; δCH$_3$ **0.49**, **0.82**	9.03	4.28	
Cys 16	7.49	4.78	2.84, 3.87		7.56	4.72	
Glu 17	8.63	4.51	2.05, 2.32	γCH$_2$ 2.11, 2.29	8.60	4.52	
Gly 18	8.52	3.68, 3.93			8.60	3.67, 3.92	
Ser 19	8.72	4.42	3.90, 4.01		8.74	4.42	
Asn 20	7.97	4.81	2.58, 3.07	δNH$_2$ 7.06, 7.55	7.96	4.80	
Val 21	9.08	3.42	1.70	γCH$_3$ 0.33, 0.33	9.08	3.40	
Cys 22	8.45	5.14	2.90, 3.21		8.40	5.14	
Gly 23	9.19	3.84, 4.06			9.21	3.83, 4.02	
Gln 24	8.50	4.07	2.05, 2.05	γCH$_2$ 2.48; εNH$_2$ 6.96, 7.67	8.54	4.05	
Gly 25	8.97	3.62, 4.26			*8.73*	3.62, 4.24	
Asn 26	8.04	5.53	2.66, 2.75	δNH$_2$ 6.12, 6.77	8.02	5.50	
Lys 27	9.30	4.45	1.20, 1.27	γCH$_2$ 0.98, 1.53; δCH$_2$ 1.36, 1.43; εCH$_2$ 2.67, 2.80; ζNH$_3^+$ 7.64	9.28	4.43	
Cys 28	9.08	5.37	2.87, 2.98		9.07	5.38	
Ile 29	9.89	4.30	2.05	γCH$_2$ 1.41, 1.53; γCH$_3$ 0.87; δCH$_3$ 0.77	9.85	4.30	
Leu 30	8.46	4.20	1.58, 1.58	γH 1.44; δCH$_3$ 0.62, 0.62	8.46	4.17	
Gly 31	8.97	3.53, 4.05			8.99	3.52, 4.03	
Ser 32	8.31	4.66	3.80, 3.80		*8.20*	4.63	
Asp 33	8.96	4.41	2.74, 2.95		*9.07*	4.45	βCH$_2$ 2.83, *3.11*
Gly 34	8.62	3.65, 4.23			8.64	3.63, 4.20	
Glu 35	7.82	4.41	2.13, 2.19	γCH$_2$ 2.45, 2.55	*7.67*	4.40	γCH$_2$ *2.66*, –
Lys 36	8.36	4.32	1.76, 2.00	γCH$_2$, 1.52, 1.52; δCH$_2$ 1.71, 1.71; εCH$_2$ 3.01, 3.01; ζNH$_3^+$ 7.59	8.33	4.29	
Asn 37	8.60	4.82	2.71, 2.71	δNH$_2$ 7.10, 7.91	8.53	4.79	
Gln 38	8.47	4.76	1.78, 1.94	γCH$_2$ 2.20, 2.31; εNH$_2$ 6.79, 7.52	8.48	*4.63*	
Cys 39	9.99	5.45	2.80, 3.16		*9.11*	5.46	
Val 40	9.26	4.90	2.35	γCH$_3$ 0.85, 0.91	9.29	4.87	
Thr 41	8.30	4.47	4.05	γCH$_3$ 1.29	8.30	4.46	
Gly 42	8.25	3.66, 4.16			8.25	3.66, 4.10	
Glu 43	8.13	4.16	1.96, 1.96	γCH$_2$ 2.37, 2.51	8.14	4.21	γCH$_2$ 2.60, –
Gly 44	7.80	3.53, 4.59			7.78	3.53, 4.50	
Thr 45	8.17	4.84	3.98	γCH$_3$ 1.23	8.20		
Pro 46		4.66	1.90, 2.25	γCH$_2$ 2.10, 2.10; δCH$_2$ 3.76, 3.93		4.60	
Lys 47	8.38	4.41	1.68, 1.72	γCH$_2$ 1.50, 1.55; δCH$_2$ 1.66, 1.74; εCH$_2$ 3.01, 3.17	8.42	4.43	
Pro 48		4.46	1.90, 2.34	γCH$_2$ 2.01, 2.06; δCH$_2$ 3.63, 4.02		4.34	
Gln 49	8.64	4.34	1.97, 2.14	γCH$_2$ 2.34, 2.44; εNH$_2$ 6.92, 7.73	8.61	4.34	
Ser 50	8.41	4.41	3.83, 3.83		8.39	4.42	
His 51	8.70	4.74	3.18, 3.28	δ^{2}H 7.31; ε^{1}H 8.60	8.69	4.71	
Asn 52	8.60	4.75	2.73, 2.83	δNH$_2$ 6.94, 7.65	8.54	4.69	
Asp 53	8.51	4.63	2.72, 2.80		**8.48**	**4.63**	βCH$_2$ ***2.83***, ***2.92***
Gly 54	8.38	3.91, 3.93			**8.38**	**3.90**, **3.90**	
Asp 55	8.12	4.63	2.62, 2.72		**8.16**	**4.66**	βCH$_2$ ***2.73***, ***2.83***
Phe 56	8.18	4.60	3.03, 3.14	δH 7.23; εH + ζH 7.34	8.18	4.57	
Glu 57	8.19	4.30	1.91, 2.00	γCH$_2$ 2.31, 2.31	8.16	4.30	
Glu 58	8.30	4.33	1.94, 2.03	γCH$_2$ 2.37, 2.37	8.25	4.33	
Ile 59	8.24	4.46	1.82	γCH$_2$ 1.17, 1.50; γCH$_3$ 0.97; δCH$_3$ 0.86	8.18	4.44	
Pro 60		4.39	1.88, 2.31	γCH$_2$ 1.98, 2.04; δCH$_2$ 3.68, 3.89		4.37	
Glu 61	8.45	4.21	1.97, 1.97	γCH$_2$ 2.41, 2.41	8.36	4.23	
Glu 62	8.39	4.23	1.91, 1.95	γCH$_2$ 2.16, 2.26	8.30	4.25	
Tyr 63	8.15	4.62	2.93, 3.11	δH 7.12; εH 6.81	8.11	4.60	
Leu 64	8.07	4.37	1.61, 1.61	γH 1.54; δCH$_3$ 0.86, 0.92	8.03	4.34	
Gln 65	7.88	4.19	1.95, 2.13	γCH$_2$ 2.30, –	*8.12*	*4.30*	

[a] The protons with chemical shift differences between natural hirudin [data from Sukumaran et al. (1987)] and recombinant desulfatohirudin greater than 0.1 ppm are boldfaced. The protons, which showed a pH-dependent chemical shift change of 0.05–0.1 ppm between pH 4.5 and pH 3.0 are underlined; those with a chemical shift change in excess of 0.1 ppm are italicized.

residues. Disulfide bridge formation between Cys 28 and Cys 39 could be excluded because these two cysteines are located on opposite sides of the same β-sheet (Figure 5B) (Richardson, 1981). The two remaining possible combinations, Cys 16–Cys

Table II: Stereospecific ^{1}H NMR Assignments in Recombinant Desulfatohirudin

residue	method[a]	resonance assignment (ppm)	
Val 2	B	$C^{\gamma 1}H_3$ 0.82	$C^{\gamma 2}H_3$ 0.85
Tyr 3	H	$H^{\beta 2}$ 2.65	$H^{\beta 3}$ 2.79
Asp 5	H	$H^{\beta 2}$ 2.62	$H^{\beta 3}$ 2.77
Cys 6	H	$H^{\beta 2}$ 2.68	$H^{\beta 3}$ 3.18
Glu 8	H	$H^{\beta 2}$ 1.75	$H^{\beta 3}$ 2.05
Ser 9	H	$H^{\beta 2}$ 4.04	$H^{\beta 3}$ 3.95
Gly 10	B	$H^{\alpha 1}$ 3.32	$H^{\alpha 2}$ 4.61
Gln 11	A	$H^{\epsilon 1}$ 6.73	$H^{\epsilon 2}$ 7.82
Asn 12	A	$H^{\delta 1}$ 6.12	$H^{\delta 2}$ 7.48
Cys 14	H	$H^{\beta 2}$ 3.39	$H^{\beta 3}$ 3.13
Leu 15	B	$C^{\delta 1}H_3$ 0.82	$C^{\delta 2}H_3$ 0.49
Cys 16	H	$H^{\beta 2}$ 3.87	$H^{\beta 3}$ 2.84
Asn 20	H	$H^{\beta 2}$ 3.07	$H^{\beta 3}$ 2.58
	A	$H^{\delta 1}$ 7.06	$H^{\delta 2}$ 7.55
Cys 22	H	$H^{\beta 2}$ 3.21	$H^{\beta 3}$ 2.90
Gln 24	A	$H^{\epsilon 1}$ 6.95	$H^{\epsilon 2}$ 7.65
Asn 26	H	$H^{\beta 2}$ 2.66	$H^{\beta 3}$ 2.75
	A	$H^{\delta 1}$ 6.77	$H^{\delta 2}$ 6.11
Cys 28	H	$H^{\beta 2}$ 2.87	$H^{\beta 3}$ 2.98
Glu 35	H	$H^{\beta 2}$ 2.20	$H^{\beta 3}$ 2.11
Asn 37	A	$H^{\delta 1}$ 7.10	$H^{\delta 2}$ 7.91
Gln 38	A	$H^{\epsilon 1}$ 6.79	$H^{\epsilon 2}$ 7.52
Cys 39	H	$H^{\beta 2}$ 2.80	$H^{\beta 3}$ 3.16
Val 40	B	$C^{\gamma 1}H_3$ 0.90	$C^{\gamma 2}H_3$ 0.85
Pro 46	H	$H^{\beta 2}$ 1.90	$H^{\beta 3}$ 2.25
Pro 48	H	$H^{\beta 2}$ 1.90	$H^{\beta 3}$ 2.34
Gln 49	A	$H^{\epsilon 1}$ 6.92	$H^{\epsilon 2}$ 7.73
Asn 52	A	$H^{\delta 1}$ 6.93	$H^{\delta 2}$ 7.63
Phe 56	H	$H^{\beta 2}$ 3.03	$H^{\beta 3}$ 3.14
Pro 60	H	$H^{\beta 2}$ 1.88	$H^{\beta 3}$ 2.31

[a] A: The individual assignment of the side-chain amide protons in Asn and Gln was based on intraresidual NOEs with βCH_2 or γCH_2, respectively. B: Assignment by reference to the low-resolution tertiary structure. H: Assignment derived with the program HABAS.

Table III: Stereospecific Assignment for Leu 15 $C^{\delta}H_3$ Obtained by Comparison of NOE Intensities with the Corresponding Distances in the Desulfatohirudin Structures Computed from a Preliminary Set of NMR Data

	NOE intensity[a]		range of the distance difference[b]	stereospecific assignment[c]	
proton X	δ = 0.82	δ = 0.49		A	B
Leu 15 HN	3	10	0.8–1.3	*	
Leu 15 Hα	5	8	1.1–1.6	*	
Cys 14 Hα	0	2	1.5–2.4	*	
Cys 16 HN	0	3	0.6–1.3	*	
Val 21 Hα	1	9	0.8–2.1	*	
Val 21 Hβ	0	3	0.9–3.1	*	
Cys 22 HN	0	4	1.5–2.2	*	

[a] Relative intensities (in arbitrary units) of the two NOEs from proton X to the δ-methyl protons of Leu 15 observed at 0.82 and 0.49 ppm. [b] The distance difference is defined as $d(X,C^{\delta 1}H_3) - d(X,C^{\delta 2}H_3)$. The range of the distance differences in five low-resolution DISMAN solutions is shown (see text). [c] A: $\delta(C^{\delta 1}H_3)$ = 0.82 ppm and $\delta(^{\delta 2}H_3)$ = 0.49 ppm were assumed. B: Alternative assignment. For each NOE the stereospecific assignment consistent with the corresponding ^1H–^1H distances is indicated by an asterisk (*).

28 and Cys 22–Cys 39 or Cys 16–Cys 39 and Cys 22–Cys 28, were distinguished from DISMAN calculations with the polypeptide segment 1–49 of desulfatohirudin. The results of two series of calculations were compared, which used input data that were identical except for these two different possible disulfide combinations. With the disulfide bridges at the same positions as in natural hirudin, four calculations converged to values of the target function between 2.6 and 3.4. With the alternative locations of the disulfide bridges, the target function values of the best four structures were in the range from 15 to 30, and in each of the four structures serious local constraint

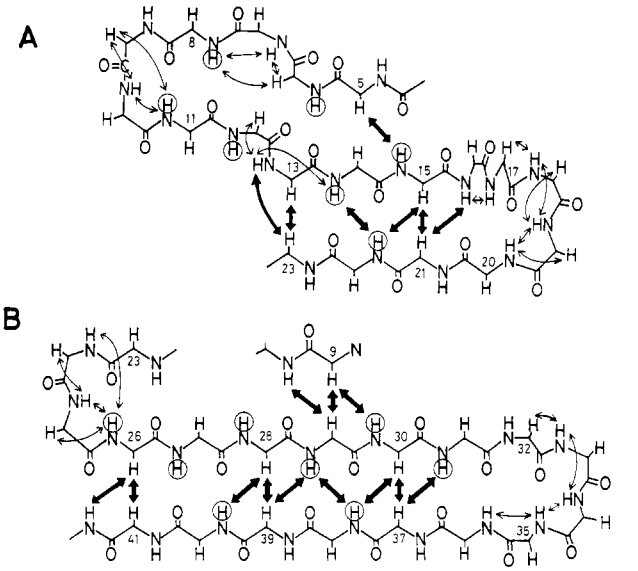

FIGURE 5: Schematic representation of the second structure elements in recombinant desulfatohirudin, which were identified prior to the distance geometry calculations (see text). Interstrand long-range backbone NOEs in the β-sheets are indicated by thick arrows. Thin arrows indicate sequential and medium-range NOEs which were essential for the characterization of the turn regions. The slowly exchanging, hydrogen-bonded amide protons are circled.

Table IV: Hydrogen Bonds Identified in Recombinant Desulfatohirudin

donor[a]	acceptor	score[b]
Cys 6 HN	Leu 15 O	25
*Gly 10 HN	Cys 28 O	36
*Gln 11 HN	Glu 8 O	36
Asn 12 HN	Thr 45 O	24
Asn 12 $HN^{\delta 2}$	Asn 26 O	28
Leu 13 HN	Cys 22 O	21
*Cys 14 HN	Cys 22 O	36
Leu 15 HN	Thr 4 O	36
*Cys 16 HN	Asn 20 O	36
*Asn 20 HN	Glu 17 O	36
ΔAsn 20 $HN^{\delta 2}$	Glu 17 O	26
*Cys 22 HN	Cys 14 O	36
*Asn 26 HN	Gly 23 O	36
*Lys 27 HN	Val 40 O	36
Cys 28 HN	Gln 11 O	36
*Ile 29 HN	Gln 38 O	35
Leu 30 HN	Ser 9 O^{γ}	18
*Gly 31 HN	Lys 36 O	30
ΔSer 32 HN	Glu 35 O	35
Glu 35 HN	Asp 33 O	20
*Gln 38 HN	Ile 29 O	36
*Val 40 HN	Lys 27 O	34
*Gly 42 HN	Gly 25 O	36
ΔThr 45 HN	Gly 10 O	24

[a] Hydrogen bonds indicated by asterisks (*) were included in the input for the DISMAN calculations. Those indicated by triangles (Δ) are implicated by the molecular geometry of the DISMAN solutions, but the exchange of the amide proton is rapid. [b] The hydrogen bonds were characterized by the hydrogen–oxygen distance (d < 2.4 Å → 2 points, 2.4 Å < d < 3.2 Å → 1 point) and the angle between the bond from the hydrogen atom to the donor and the line connecting the donor with the acceptor (α < 35° → 2 points, 35° < α < 70° → 1 point). For each potential hydrogen bond, the points thus scored in the DISMAN structures I–IX were summed up. All hydrogen bonds scoring higher than 18 points are listed.

violations were observed around the assumed disulfide bridges Cys 22–Cys 28 and Cys 16–Cys 39. For example, the violation of the distance constraint S^{γ}(C22)–S^{γ}(C28) was always greater than 0.7 Å, and the violation of the χ^1 angle of Cys 28 was in the range from 26° to 55°. Constraints enforcing the disulfide bonds 6–14, 16–28, and 22–39 (see Materials and

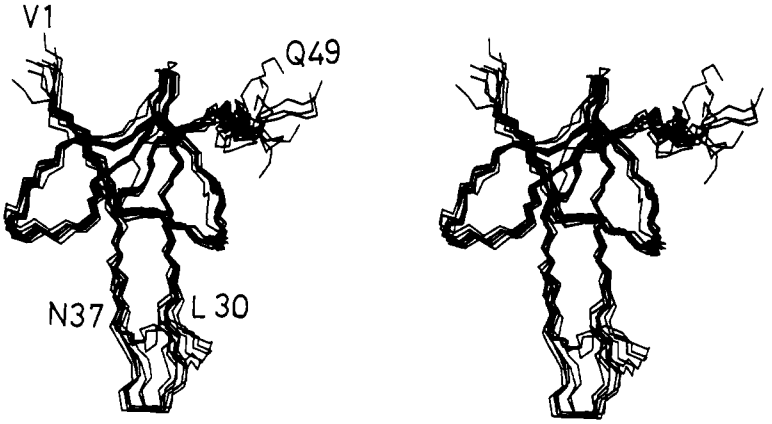

FIGURE 6: Stereoview of the polypeptide chain from residue 1 to residue 49 in the nine DISMAN structures of recombinant desulfatohirudin. The nine conformers were superimposed for pairwise minimum RMSD of the backbone heavy atoms N, C^{α}, and C′ with structure I, whereby only the polypeptide segments from Tyr 3 to Leu 30 and from Asn 37 to Pro 48 were used for the calculation of the RMSD.

Table V: Residual Constraint Violations for the Polypeptide Segment 1–49 in the Nine DISMAN Solutions of Recombinant Desulfatohirudin

upper bounds[a]	av violation (Å)	no. of violations[b] 0.1–0.3 Å	0.3–0.5 Å
I	0.02	18	5 (0.48)
II	0.02	30	5 (0.46)
III	0.02	29	7 (0.45)
IV	0.02	27	6 (0.51)
V	0.02	21	8 (0.55)
VI	0.02	27	4 (0.39)
VII	0.02	28	7 (0.48)
VIII	0.02	38	5 (0.45)
IX	0.03	35	6 (0.43)

lower bounds[c]	av violation (Å)	no. of violations 0.1–0.4 Å[b]	dihedral angles > 5 deg[d]
I	0.04	6 (0.26)	1 (5)
II	0.05	11 (0.20)	3 (12)
III	0.04	7 (0.20)	2 (7)
IV	0.05	6 (0.35)	5 (7)
V	0.05	9 (0.28)	2 (8)
VI	0.02	7 (0.14)	3 (7)
VII	0.04	6 (0.27)	4 (8)
VIII	0.04	9 (0.29)	3 (7)
IX	0.04	10 (0.20)	4 (7)

[a] Upper bounds from NOEs, hydrogen bonds, and disulfide bonds. [b] These columns list the number of distance constraint violations in the ranges indicated. In parentheses, the maximum violation is indicated. [c] Lower bounds from hydrogen bonds, disulfide bonds, and van der Waals volumes. The violations of van der Waals constraints were checked against the values of the repulsive core radii as given in Table I of Braun and Gō (1985). [d] The number of violations of dihedral angle constraints larger than 5 deg is listed. In parentheses, the maximum violation is indicated.

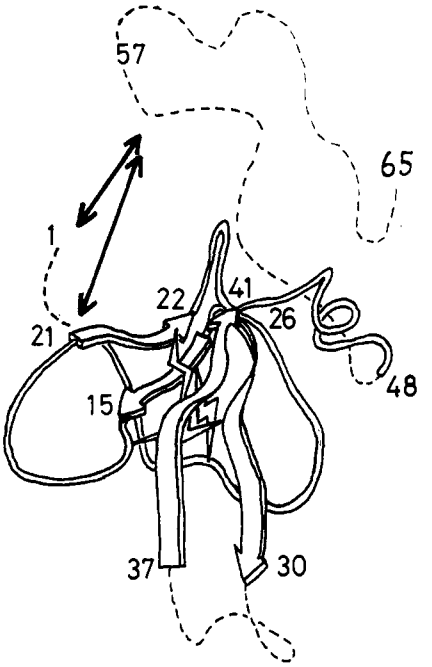

FIGURE 7: Schematic ribbon drawing of recombinant desulfatohirudin generated with the program CONFOR (Billeter et al., 1985a,b). The coordinates of DISMAN solution I were used, and the orientation is the same as in Figure 6. The arrowed ribbons indicate the position and direction of the β-sheet strands. A rope structure was used for the loop areas, with the broken lines indicating those segments which are not well constrained by the NMR data. The disulfide bridges are indicated by lightning bolts. The arrows indicate the positions of NOEs between the protein core and the flexible C-terminal segment 49–65.

Methods) were therefore used in all subsequent calculations.

Structure Calculations with the Program DISMAN. Figure 3 shows that NOEs relating the C-terminal polypeptide segment 49–65 with the core of the protein were observed only between Phe 56, Val 1, and Val 21. Therefore, two independent calculations were carried out, one for the whole molecule with residues 1–65 using the complete input data set and the other one for the polypeptide segment 1–49 using only the input data relating to residue pairs in this segment. With 50 starting structures, five and four calculations converged at levels 65 (structures I–V), and 49 (structures VI–IX), respectively. A survey of the residual constraint violations is given in Table V for the segment 1–49 in all nine structures. The violations are classified according to the type of constraint violated and the size of the violation. The nine structures show similar levels of average violations. In each structure, all NOE violations are smaller than 0.5 Å, except for three violations in structures IV and V, which are slightly higher than 0.5 Å. The number of upper-limit violations greater than 0.3 Å is in the range from four to eight per DISMAN solution, which is less than 2% of the total number of upper-limit constraints. The size of the lower-limit and van der Waals constraint violations is less than 0.3 Å, with a single exception in structure IV. The number of dihedral angle violations exceeding 5° is from one to five per structure, which is less than 4% of the total number of dihedral angle constraints.

In Figure 6 the polypeptide backbone segment from 1 to 49 of all nine structures is superimposed for minimal RMSD. The C-terminal part beyond Gln 49 is not shown, since the individual DISMAN solutions obtained from the calculations with hirudin 1–65 deviate significantly in this region. The

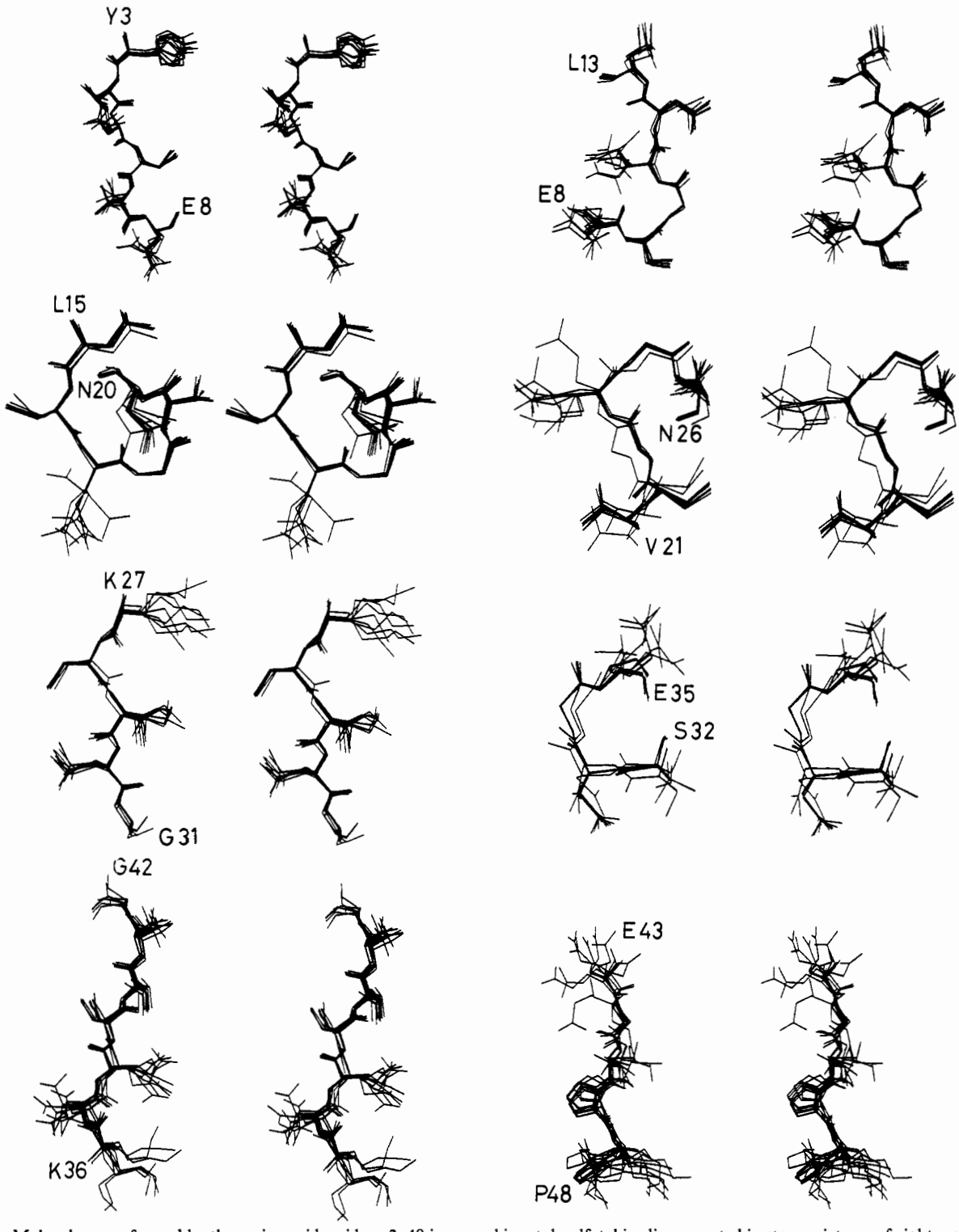

FIGURE 8: Molecular core formed by the amino acid residues 3–48 in recombinant desulfatohirudin presented in stereo pictures of eight segments. The individual segments from the nine DISMAN solutions were superimposed for a best local fit with structure I, where the backbone atoms N, C^{α}, and C′ over the entire length of the segments were considered for the fitting procedure. All heavy atoms have been plotted.

situation is schematically shown in Figure 7, where the two NOE constraints observed between protons in the segments 1–49 and 50–65 are indicated by arrows (see also Figure 3). Figure 8 shows the complete structure 3–48 with side chains, presented as stereoviews of eight peptide segments superimposed for local minimum RMSD of the backbone heavy atoms.

DISCUSSION

A first criterion for the evaluation of the structures obtained for recombinant desulfatohirudin is provided by the statistics on the residual violations of the input constraints (Table V). On the basis of experience with other proteins [e.g., Wagner et al. (1987), Steinmetz et al. (1987), and Kline et al. (1988)], we conclude that all of the structures I–IX in Table V represent satisfactory solutions to the geometric problem of finding a three-dimensional polypeptide fold that is compatible with the experimental data. This judgement is based both on the average of the residual violations and on the fact that there are no outstandingly large residual violations of individual constraints. In calculations I–V good fits were obtained also for the constraints involving protons of residues 50–65, which is not surprising in view of the scarcity of data for this peptide

Table VI: RMSD Values (Å) for Pairwise Global Comparisons of the Nine DISMAN Solutions I–IX of Recombinant Desulfatohirudin[a]

	I	II	III	IV	V	VI	VII	VIII	IX
I		0.56	0.65	0.71	0.78	0.76	0.69	1.17	0.79
II	1.61		0.48	0.47	0.79	0.70	0.75	1.00	0.90
III	1.75	1.43		0.54	0.78	0.71	0.62	0.90	0.94
IV	1.67	1.23	1.48		0.64	0.72	0.75	0.89	1.01
V	1.41	1.66	1.67	1.57		0.79	0.67	1.02	0.95
VI	1.58	1.69	1.89	1.72	1.68		0.86	0.93	1.02
VII	1.74	1.62	1.47	1.52	1.62	1.85		1.05	0.91
VIII	2.10	1.66	1.72	1.64	1.88	2.00	1.74		1.31
IX	1.63	1.72	1.96	1.90	1.85	2.01	1.95	2.14	

[a] Residues 3–30 and 37–48 were considered in the calculation of the RMSDs (see text). The upper right triangle lists the RMSDs calculated for the backbone atoms (N, C^{α}, C′); the lower left triangle lists the RMSDs calculated for all heavy atoms. The average of the pairwise RMSDs for the backbone atoms is 0.81 ± 0.19 Å, and the average of the pairwise RMSDs for all heavy atoms is 1.71 ± 0.20 Å.

segment (Figure 3 and 4) (Wüthrich, 1986).

A second, critical criterion for evaluating the results is the uniqueness of the structure obtained. As indicated in Figure 7, neither the conformation of the C-terminal polypeptide segment 49–65 nor its orientation relative to the other parts of the molecule is defined by the NMR data, and the structure of the N-terminal dipeptide segment as well as the loop formed by residues 31–36 is also uncertain. In contrast, for the molecular region of residues 3–30 and 37–48, the pairwise RMSDs for the different DISMAN solutions are between 0.47 and 1.31 Å for the backbone atoms and between 1.47 and 2.14 Å for all heavy atoms, with average RMSDs of 0.81 and 1.71 Å, respectively. These values are similar to those obtained for high-resolution solution structures of other small globular proteins [e.g., Kline et al. (1988)] and show that a core region in the structure of recombinant desulfatohirudin is well defined by the NMR data. When the structures calculated with respectively residues 1–65 (structures I–V) or 1–49 (structures VI–IX) are compared, there appears to be a trend for closer fit among the conformers obtained with the complete polypeptide chain. This can be traced to the fact that residues 45–49 are more variable in the structures computed with the polypeptide fragment 1–49. Otherwise, the differences are hardly significant, as can be seen by comparing the averages of the RMSDs within each of the groups I–V and VI–IX and between conformers in the two groups.

In the presentation of the molecular core in Figures 6 and 7, a double-stranded β-sheet consisting of residues Asn 26 to Glu 31 and Lys 36 to Thr 41 is in the foreground, and a short β-sheet consisting of residues Leu 13 to Cys 16 and Val 21 to Gly 23 is in the back. The sequence locations of the two β-sheets coincide exactly with those found by the empirical pattern recognition approach (Figure 5). The relative spatial locations of the two β-sheets are fixed by the two disulfide bridges Cys 16–Cys 28 and Cys 22–Cys 39. A polypeptide loop closed by the disulfide bridge Cys 6–Cys 14 is located on the right side of the two β-sheets. The spatial arrangement of this loop relative to the β-sheets is defined by three hydrogen bonds, Cys 6(NH)–Leu 15(O′), Gly 10(NH)–Cys 28(O′), and Cys 28(NH)–Gln 11(O′).

A complete description of the desulfatohirudin core including the amino acid side chains is afforded by Figure 8 and Table VII. The variation of the individual dihedral angles among the nine DISMAN solutions (Table VII) represents a basis for evaluating the precision with which the local conformations of individual parts of the polypeptide chain are defined. In the segment 3–8 (Figure 8) the first four residues are well defined, except for the orientation of the aromatic ring of Tyr 3 about the dihedral angle χ^2. This disordered state of the ring occurs in spite of the large number of NOE constraints involving this side chain (Figure 4). It originates mainly from the use of pseudoatoms for the ring protons, which was inevitable because of the chemical shift degeneracy of the symmetry-related ring protons (Wüthrich et al., 1983). For residues Thr 7 and Glu 8 the side chains beyond C^{β} are largely

Table VII: Individual Dihedral Angles (deg) and Their Variations among the Nine DISMAN Structures of Recombinant Desulfatohirudin[a]

residue	Φ	ψ	χ^1	χ^2
Val 1	#	119 ± 28	#	
Val 2	−87 ± 24	112 ± 22	160 ± 5	
Tyr 3	−110 ± 17	166 ± 5	−90 ± 5	−42 ± 91
Thr 4	−120 ± 7	175 ± 5	55 ± 9	
Asp 5	−74 ± 5	151 ± 8	−66 ± 21	
Cys 6	−61 ± 13	169 ± 5	−63 ± 20	
Thr 7	−127 ± 5	5 ± 5	#	
Glu 8	−161 ± 5	143 ± 5	−89 ± 8	#
Ser 9	−66 ± 5	131 ± 8	150 ± 5	
Gly 10	92 ± 5	3 ± 12		
Gln 11	−76 ± 11	156 ± 11	−138 ± 32	#
Asn 12	−124 ± 7	−178 ± 5	−54 ± 32	
Leu 13	51 ± 10	56 ± 5	−56 ± 5	122 ± 9
Cys 14	−168 ± 5	156 ± 8	38 ± 7	
Leu 15	−73 ± 5	153 ± 9	−46 ± 5	−164 ± 6
Cys 16	−120 ± 6	−88 ± 5	−92 ± 10	
Glu 17	−116 ± 8	160 ± 6	#	−160 ± 49
Gly 18	51 ± 5	17 ± 38		
Ser 19	28 ± 76	42 ± 5	−178 ± 74	
Asn 20	−109 ± 5	143 ± 5	−69 ± 11	
Val 21	−69 ± 5	144 ± 4	135 ± 9	
Cys 22	−139 ± 7	140 ± 6	−169 ± 7	
Gly 23	136 ± 7	−154 ± 5		
Gln 24	35 ± 5	90 ± 7	−60 ± 60	#
Gly 25	101 ± 7	−36 ± 5		
Ile 29	−118 ± 5	171 ± 5	−28 ± 5	
Lys 27	−121 ± 5	143 ± 10	−56 ± 93	164 ± 56
Cys 28	−91 ± 5	114 ± 16	−173 ± 1	
Ile 29	−87 ± 14	150 ± 10	−122 ± 26	#
Leu 30	−136 ± 10	129 ± 16	−132 ± 10	98 ± 14
Gly 31	−66 ± 8	−132 ± 27		
Ser 32	168 ± 40	82 ± 17	#	
Asp 33	97 ± 32	70 ± 28	−48 ± 29	
Gly 34	93 ± 17	−30 ± 22		
Glu 35	−91 ± 22	158 ± 9	178 ± 5	#
Lys 36	179 ± 89	166 ± 24	−53 ± 18	#
Asn 37	−105 ± 35	141 ± 12	#	
Gln 38	−147 ± 10	158 ± 27	−104 ± 43	#
Cys 39	−126 ± 20	101 ± 7	−167 ± 5	
Val 40	−106 ± 9	150 ± 5	−41 ± 5	
Thr 41	−79 ± 9	156 ± 7	#	
Gly 42	132 ± 16	#		
Glu 43	#	123 ± 30	#	#
Gly 44	−114 ± 37	156 ± 53		
Thr 45	−94 ± 58	138 ± 20	#	
Pro 46	146 ± 16			
Lys 47	−50 ± 14	#	#	#
Pro 48	#			
Gln 49	−57 ± 17	−20 ± 62	#	#

[a] For each angle the range of values covered by the nine DISMAN structures is represented by the center of the range and its size. Angles relating to methyl and NH_3^+ groups have not been listed. The symbol # indicates that the torsion angle range is larger than 180°.

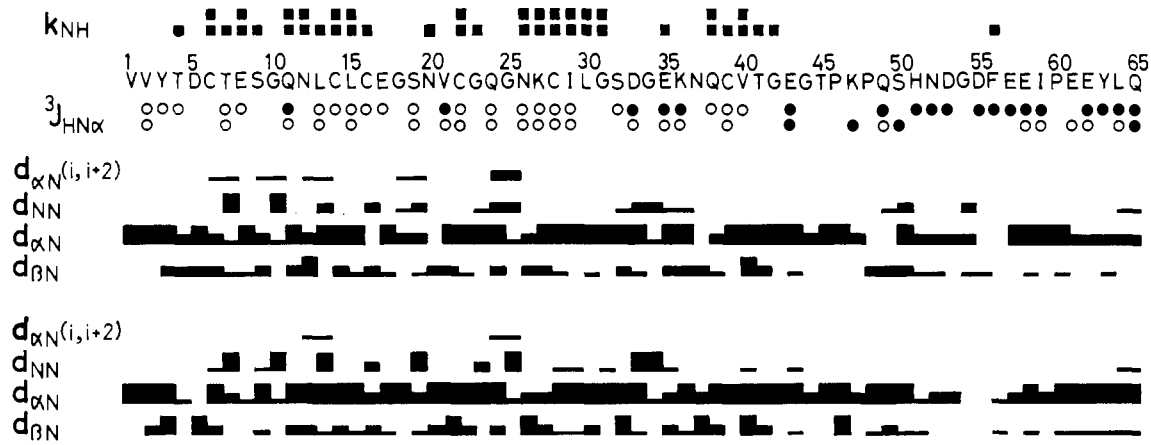

FIGURE 9: Survey of the experimental data (except for the long-range, interstrand NOEs shown in Figure 5) which characterize the secondary structure in recombinant desulfatohirudin. Also included are, directly below the corresponding desulfatohirudin data, the data for natural hirudin [from Sukumaran et al. (1987)]. Above the amino acid sequence, filled squares identify those residues in desulfatohirudin for which the amide proton exchange is sufficiently slow, so that the NH resonances could be observed in a NOESY experiment started immediately after dissolving the protein in ^2H_2O at pH 4.6 and 22 °C and completed in 6 h. For natural hirudin the corresponding data represent slowly exchanging amide protons still present 24 h after dissolving the protein in ^2H_2O at pH 3.0 and 25 °C (Sukumaran et al., 1987). Below the sequence, open circles indicate residues with $^3J_{HN\alpha} > 8.0$ Hz and filled circles those with $^3J_{HN\alpha} < 8.0$ Hz. At the bottom, the thickness of the bars used to represent sequential and medium-range NOEs reflects the intensity of the NOESY cross peaks.

disordered. Residues 5–8 form a quite typical half-turn, which could also be recognized directly from the pattern of local NOE constraints and spin–spin coupling constants (Wagner et al., 1986; Wüthrich, 1986). In the segment 9–13 all residues are well defined, except for the peripheral parts of Gln 11. The segment 8–11 forms a quite typical type II tight turn. In the following segment 15–20, local disorder prevails at Ser 19 and for the side chain of Glu 17. In the segment 21–26, one of the nine DISMAN solutions is different from the others over the full length, but except for the side chain of Gln 24, the other eight solutions form a tight group. As can be seen in Figure 8 the two tetrapeptide segments 17–20 and 23–26 form tight turns, but the dihedral angles (Table VII) do not correspond to any of the classical turns. The segment 27–31 is overall well ordered, except for the side chains of Lys 27 and Ile 29. For one of the DISMAN solutions, the plane of the peptide bond 28–29 is rotated away from the orientation in the other eight conformers. Residues 32–35 form a type I′ turn. In spite of the different global orientations of this loop in two groups of structures (Figure 6), the local fit is quite good. Nonetheless, most of the backbone dihedral angles of residues 31–37 show relatively large variability (Table VII), so that the appearance of two groups of conformers cannot be entirely attributed to a localized hinge motion at residues 31 and 36. In the segment 36–42 the side chains of the first three residues are noticeably disordered, but residues 40–42 are well defined. In the final segment of the protein core, 43–48, all the side chains are only poorly constrained. Comparison with Figure 4 shows that there is a trend to increased disorder in the regions with few NOE constraints. Examples are the segments 31–34 and 40–48, but there are also individual disordered side chains that have numerous NOE constraints. A more clear-cut correlation was found with the stereospecific assignments (Table II); i.e., when stereospecific assignments could be obtained, the side-chain conformation is generally well confined by the NMR data.

As an additional structural detail, the hydrogen bonds in desulfatohirudin 1–49 were identified with the same criteria as used previously for Tendamistat (Kline et al., 1988). These criteria are described in detail in the footnote to Table IV. In addition to the 13 hydrogen bonds used in the input for the DISMAN calculations (Figure 5, Table IV), 11 hydrogen bonds were thus identified in the three-dimensional structure. For eight of these, the hydrogen-bond formation indicated by the molecular geometry is also supported by slow exchange of the amide proton. For the remaining three amide protons, slowed exchange was not observed, which is probably related to their location near the protein surface.

Comparison of the presently determined structure of recombinant desulfatohirudin with the structure presented for natural hirudin in solution (Sukumaran et al., 1987; Clore et al., 1987) shows that the global arrangement of residues 3–30 and 37–48 in the well-defined protein core is nearly identical in the two molecules. Furthermore, in both molecules the C-terminal segment 49–65 appears to be largely disordered, and the loop including residues 31–36 is not uniquely defined. The close coincidence of the polypeptide backbone structures in the two molecules is convincingly documented by a comparison of the raw NMR data used to determine the secondary structure (Figure 5), i.e., the sequential and medium-range NOE connectivities, $^3J_{HN\alpha}$, and the positions of slowly exchanging amide protons (Figure 9). For the short-range NOEs the similarity between natural hirudin and recombinant desulfatohirudin is very high. Most differences seen in Figure 9 consist of corresponding NOEs being reported as "weak" or "strong". For the spin–spin couplings $^3J_{HN\alpha}$ there is also extensive agreement. Here, a more complete set of data was collected for the recombinant protein, and in the data for natural hirudin there appears to be a tendency for systematic overestimation by about 2 Hz of spin–spin couplings with actual values around 6–8 Hz (Wüthrich, 1986). Finally, all slowly exchanging amide protons in desulfatohirudin were also observed in the natural protein. For the latter, some additional slowly exchanging amide protons were reported (Figure 9), which must be due to the different experimental conditions used [see also Haruyama et al. (1989)]. Considering the data compiled in Figure 9, it is not surprising that the architecture of the regular secondary structure elements in natural hirudin (Sukumaran et al. 1987) is virtually identical with that shown in Figure 5 for recombinant desulfatohirudin.

There are certain differences in the description of local structural features in natural hirudin and desulfatohirudin, such as the identification of individual tight turns with classical turn types. Furthermore, Clore et al. (1987) describe the side chains of Lys 27, Ile 29, Gln 38, Thr 41, and Lys 47, which all are quite outstandingly disordered in desulfatohirudin

(Table VII, Figure 8), as "relatively well defined", as compared to most of the side chains in their structure being "rather poorly defined". From the available statistics on the two structure determinations it is, however, clear that the structure of the natural hirudin was significantly less well characterized. This is a consequence of the smaller number of conformational constraints obtained and the absence of stereospecific assignments (Güntert et al., 1989). Therefore, we do not feel that the forementioned apparent local structure differences are well substantiated, and with the available structural data we can really only conclude that the absence of the sulfato group at Tyr 63 does not noticeably affect the global three-dimensional structure of hirudin.

The structure of recombinant desulfatohirudin was checked in two ways against the ^{1}H NMR spectra. First, qualitative comparisons were made between outstandingly large observed conformation-dependent chemical shifts and the shifts predicted from the well-defined part of the molecular structure. There was no evidence for large ring current shifts, but several backbone proton shifts in excess of 1.0 ppm were attributed to hydrogen bonding. Calculations of these shifts from the desulfatohirudin structure using eq 2 and 3 from Wagner et al. (1983) gave satisfactory coincidence for all but two protons, the exceptions being shifts of −1.06 ppm for Leu 13 αH and 1.68 ppm for Cys 39 NH. Second, a hypothetical NOESY spectrum was computed from the five DISMAN solutions for desulfatohirudin 1–65, using the assumption that all ^{1}H–^{1}H distances shorter than 3.5 Å should be observable [see Kline et al. (1988) for a description of the procedures used]. On the one hand, the NOESY cross peaks predicted from the well-defined core of the protein coincided in all instances with experimental cross peaks. On the other hand, a much larger percentage of the experimental NOEs remained unaccounted for than in similar checks with more completely structured proteins, e.g., Tendamistat (Kline et al., 1988). The implication is that some NOESY cross peaks might be due to transient near approach of proton pairs associated with flexible parts of the molecular structure, i.e., the polypeptide segments 31–36 and 49–65, or individual amino acid side chains.

In conclusion, those molecular regions that have a well-defined three-dimensional structure in natural hirudin (Clore et al., 1987) and in recombinant desulfatohirudin (Figure 6) are very similar in the two molecules. In a search for a possible structural basis for the reduced activity of desulfatohirudin, one must, however, consider that all experimental data presented in this paper are not accounted for by the structure of Figures 6–8. It is conceivable that the activity depends on the presence of transient conformational species that are not readily observable by NOEs but could cause the forementioned, so far unexplained outstanding chemical shifts of Leu 13 αH and Cys 39 NH. Different interactions between the protein core and the flexible peptide segments in natural hirudin and desulfatohirudin could also be the cause of the chemical shift variations between the two proteins, which were observed for residues 5, 9, 15, and 53–55 (Table I). We have started further investigations of the NMR data that cannot be accounted for by the molecular conformation characterized in Figures 6–8 (Haruyama et al., 1989). This might be of interest also beyond the hirudin project, since the simultaneous presence in the same molecule of well-structured domains and flexible polypeptide segments is also indicated for other proteins in solution [e.g., Steinmetz et al. (1987)].

ACKNOWLEDGMENTS

We thank Dr. W. Märki and H. Rink of Ciba-Geigy AG. for arranging the gift of recombinant desulfatohirudin and for providing unpublished information on the chemistry of the protein, Dr. W. Braun for the use of the program DISMAN, and E. Huber for the careful processing of the manuscript. We acknowledge the use of the computing facilities at the Zentrum für Interaktives Rechnen (ZIR) of the ETH Zürich.

SUPPLEMENTARY MATERIAL AVAILABLE

Two tables showing NOE distance constraints used as input for the DISMAN calculations of the recombinant desulfatohirudin structure, and dihedral angle constraints from HABAS and from preliminary structure determinations for two turn regions in desulfatohirudin (17 pages). Ordering information is given on any current masthead page.

REFERENCES

Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.

Arseniev, A., Schultze, P., Wörgötter, E., Braun, W., Wagner, G., Vašák, M., Kägi, J. H. R., & Wüthrich, K. (1988) *J. Mol. Biol. 201*, 637–657.

Bagdy, D., Barabas, E., Graf, L., Ellebaek, T., & Magnusson, S. (1976) *Methods Enzymol. 45*, 669–678.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Billeter, M., Engeli, M., & Wüthrich, K. (1985a) *J. Mol. Graphics 3*, 79–83.

Billeter, M., Engeli, M., & Wüthrich, K. (1985b) *J. Mol. Graphics 3*, 97–98.

Braun, P. ., Dennis, S., Hofsteenge, J., & Stone, S. R. (1988) *Biochemistry 27*, 6517–6522.

Braun, W., & Gō, N. (1985) *J. Mol. Biol. 186*, 611–626.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Chang, J.-Y. (1983) *FEBS Lett. 164*, 307–313.

Clore, G. M., Sukumaran, D. K., Nilges, M., Zarbock, J., & Gronenborn, A. M. (1987) *EMBO J. 6*, 529–539.

Dodt, J., Müller, H. P., Seemüller, V., & Chang, J.-Y. (1984) *FEBS Lett. 165*, 180–184.

Dodt, J., Seemüller, V., Maschler, R., & Fritz, H. (1985) *Biol. Chem. Hoppe-Seyler 366*, 379–385.

Griesinger, C., Sørensen, O. W., & Ernst, R. R. (1985) *J. Am. Chem. Soc. 107*, 6394–6396.

Grossenbacher, H., Auden, J. A. L., Bill, K., Liersch, M., & Maerki, W. (1987) in *XIth International Congress on Thrombosis and Haemostasis*, Brussels, July 11, 1987, Abstract 34.

Güntert, P., Braun, W., Billeter, M., & Wüthrich, K. (1989) *J. Am. Chem. Soc.* (in press).

Haruyama, H., Qian, Y.-Q., & Wüthrich, K. (1989) *Biochemistry* (following paper in this issue).

Hyberts, S. G., Maerki, W., & Wagner, G. (1987) *Eur. J. Biochem. 164*, 625–635.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4545–4553.

Kline, A. D., Braun, W., & Wüthrich, K. (1988) *J. Mol. Biol. 204*, 675–724.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Markwardt, F. (1970) *Methods Enzymol. 19*, 924–932.

Markwardt, F., Hauptmann, J., Nowak, G., Klessen, C., & Walsmann, P. (1982) *Thromb. Haemostasis 47*, 226–229.

Meyhack, B., Heim, J., Rink, H., Zimmermann, W., & Maerki, W. (1987) in *XIth International Congress on Thrombosis and Haemostasis*, Brussels, July 11, 1987, Abstract 33.

Neuhaus, D., Wagner, G., Vašák, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257-273.
Otting, G., & Wüthrich, K. (1986) *J. Magn. Reson. 66*, 359-363.
Rance, M., Sørenson, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1984) *Biochem. Biophys. Res. Commun. 117*, 479-485.
Rance, M., Bodenhausen, G., Wagner, G., Wüthrich, K., & Ernst, R. R. (1985) *J. Magn. Reson. 62*, 497-510.
Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250-254.
Richardson, J. (1981) *Adv. Protein Chem. 34*, 168-339.
Senn, H., Billeter, M., & Wüthrich, K. (1984) *Eur. Biophys. J. 11*, 3-15.
Steinmetz, W. E., Bougis, P. E., Rochat, H., Redwine, O. D., Braun, W., & Wüthrich, K. (1987) *Eur. J. Biochem. 172*, 101-116.
Stone, S. R., & Hofsteenge, J. (1986) *Biochemistry 25*, 4622-4628.
Sukumaran, D. K., Clore, G. M., Preuss, A., Zarbock, J., & Gronenborn, A. M. (1987) *Biochemistry 26*, 333-338.
Wagner, G. (1983) *J. Magn. Reson. 55*, 151-156.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347-366.
Wagner, G., & Zuiderweg, E. R. P. (1983) *Biochem. Biophys. Res. Commun. 113*, 854-860.
Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948-5949.
Wagner, G., Neuhaus, D., Wörgötter, E., Kägi, J. H. R., & Wüthrich, K. (1986) *J. Mol. Biol. 187*, 131-135.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Gō, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611-639.
Wider, G., Lee, K. H., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 367-388.
Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295-315.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949-961.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715-740.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601-611.

Static and Transient Hydrogen-Bonding Interactions in Recombinant Desulfatohirudin Studied by ^{1}H Nuclear Magnetic Resonance Measurements of Amide Proton Exchange Rates and pH-Dependent Chemical Shifts†

Hideyuki Haruyama,‡ Yan-Qiu Qian, and Kurt Wüthrich*
Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule—Hönggerberg, CH-8093 Zürich, Switzerland

Received October 7, 1988; Revised Manuscript Received January 3, 1989

ABSTRACT: With proton nuclear magnetic resonance spectroscopy at 22 °C and pD 4.5, individual exchange rates in the range from 2×10^{-5} to 1×10^{-1} min^{-1} were observed for 23 amide protons in recombinant desulfatohirudin. The remaining 38 backbone amide protons exchange more rapidly than 1×10^{-1} min^{-1}. All 23 slowly exchanging protons are located in the polypeptide segment from residue 4 to residue 42, which forms a well-defined globular domain. Three different breathing modes of this molecular region are manifested in the exchange data, which appear to be correlated with the location of the three disulfide bonds. Chemical shift changes larger than 0.15 ppm between pH 2.5 and pH 5.0 arising from through-space interactions with carboxyl groups were observed for seven backbone amide protons. Two of these shifts can be explained by hydrogen bonds in the core of the protein, Gly 25 NH–Glu 43 O$^{\epsilon}$ and Ser 32 NH–Asp 33 O$^{\delta}$, and two others by intraresidual NH–O$^{\epsilon}$ interactions in Glu 61 and Glu 62. The remaining three pH shifts for Glu 35, Cys 39, and Ile 59 imply the existence of transient interactions between the molecular core and the flexible C-terminal segment 49–65, which have so far not been characterized by nuclear Overhauser effects or other conformational constraints.

The solution conformation of recombinant desulfatohirudin determined by NMR[1] contains a well-structured globular core with residues 3–30 and 37–48 and two disordered, presumably quite flexible segments, 31–36 and 49–65 (Haruyama & Wüthrich, 1989). A similar overall structure was reported for natural hirudin (Clore et al., 1987). To further characterize the combination of well-structured and disordered flexible polypeptide segments in desulfatohirudin, the present paper describes quantitative measurements of amide proton exchange rates and the pH dependence of proton chemical shifts.

† Financial support for this project was obtained from Sankyo Co., Ltd., Tokyo, Japan, and a special grant of the ETH Zürich.
‡ Present address: Analytical and Metabolic Research Laboratories, Sankyo Co., Ltd., 2-58 Hiromachi 1-chome, Shinagawa-ku, Tokyo, 140 Japan.

[1] Abbreviations: NMR, nuclear magnetic resonance; 1D and 2D, one dimensional and two dimensional; COSY, two-dimensional correlated spectroscopy; 2QF-COSY, two-quantum-filtered COSY; RELAYED-COSY, two-dimensional relayed coherence transfer spectroscopy; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; TSP, 3-(trimethylsilyl)[2,2,3,3-^{2}H$_4$]-propionate, sodium salt.

0006-2960/89/0428-4312$01.50/0